

# LGS İngilizce Tahmin Modeli: Proje Planı ve Stratejik Analiz

## 1. Vizyon ve Proje Özeti

Biz, eğitimde ezberin yerine yorumlamayı, tahminin yerine veriyi koymak istiyoruz.

LGS İngilizce Tahmin Modeli, geçmiş yılların sorularını yalnızca analiz eden bir sistem değil; geleceğin öğrencilerine daha iyi hazırlanma fırsatı sunan bir sistemin ürünüdür. Bu proje, yapay zekâ ve veri biliminin gücünü kullanarak, LGS İngilizce sorularındaki örüntüleri, temaları ve dil yapısını anlamlandırıyor.

Amaç; öğretmenlerin, öğrencilerin ve eğitimcilerin geleceğe bir adım önden bakabilmesini sağlamak. Biz, veriyi eğitimin pusulası haline getirmek istiyoruz.

Bu vizyonu hayata geçirmek için projemiz, stratejik bir temel üzerine inşa edilmiştir. İlk aşamada, projenin SMART hedefleri ve stratejik yönü (SWOT analizi) belirlenecektir. İlk aşamayla paralel olarak, projenin "ham verisi" olan geçmiş LGS soruları toplanacak ve çok katmanlı bir etiketleme sürecine tabi tutulacaktır. Bu süreçte, sorular **konularına** (örn: Teen Life, In the Kitchen), **zorluk seviyelerine** (Kolay/Orta/Zor) ve **soru tiplerine** (Diyalog, Boşluk Doldurma vb.) göre ayrılacaktır. Son olarak, bu zenginleştirilmiş veri analiz edilerek, bir tahmin modeli stratejisi oluşturulacaktır.

## 2. SMART Hedefler

Projenin başarıyla tamamlanması için belirlenen hedefler, SMART metodolojisine göre aşağıda detaylandırılmıştır.

Kategori	Detay / Açıklama
<b>S – Spesifik (Specific)</b>	Proje, LGS İngilizce dersinden çıkışlı soruları (örn: 2018-2024 arası) analiz ederek, konu, soru tipi ve dil bilgisi yapısına göre kategorize edecek ve bu verilere dayanarak istatistiksel bir tahmin modeli geliştirecektir.
<b>M – Ölçülebilir (Measurable)</b>	Modelin konu dağılımı tahmin doğruluğu, geçmiş yıl verileriyle çapraz test edildiğinde (cross-validation) %70 ve üzerinde olacaktır. Analiz sonuçları (konu dağılımı, soru tipi trendleri) tablo ve grafiklerle net bir şekilde raporlanacaktır.
<b>A – Ulaşılabilir (Achievable)</b>	Proje, mevcut ekip yetkinlikleri (veri analizi, konu uzmanlığı) ve erişilebilir (çıkışlı sorular) veri seti ile 12 haftalık bir zaman diliminde tamamlanabilir durumdadır.
<b>R – İlgili (Relevant)</b>	Hedef, öğrencilerin ve öğretmenlerin "Hangi konulara odaklanmalıyız?" sorusuna veriye dayalı yanıtlar sunarak LGS hazırlık sürecini daha verimli hale getirme ana amacıyla doğrudan ilgilidir.
<b>T – Zaman Sınırlı (Time-bound)</b>	Proje, 12 haftalık net bir zaman planına sahiptir. (Detaylar "Proje Zaman Çizelgesi" bölümünde belirtilmiştir.)

### 3. SWOT Analizi

Projenin stratejik konumlandırmasını belirlemek amacıyla hazırlanan Güçlü Yönler, Zayıf Yönler, Fırsatlar ve Tehditler analizi aşağıdadır.

İçsel Faktörler	Dışsal Faktörler
<b>Güçlü Yönler (Strengths)</b>	<b>Fırsatlar (Opportunities)</b>
Konuya ilgili ve analistik yetkinliğe sahip ekip.  Proje amacı, hedefleri ve görev dağılımının net olması.  Tüm LGS sorularına (veri kaynağı) ücretsiz ve kolay erişim.  Veri analizi ve temel yapay zeka modelleri için gerekli teknolojiye erişim.	Tüm geçmiş yıl verilerinin internette açık kaynak olarak bulunması.  Temel istatistik ve trend analizlerinin öğrenciler için yüksek değer taşıması.  Yapay zeka araçlarının, basit tahmin modellerini hızla eğitmek için kullanılabilmesi.  Projenin, gelecekte diğer LGS dersleri için de bir prototip oluşturma potansiyeli.
<b>Zayıf Yönler (Weaknesses)</b>	<b>Tehditler (Threats)</b>
Gelişmiş makine öğrenimi ve model eğitimi konusunda derinlemesine teknik deneyim eksikliği.  Veri etiketleme sürecinin (örn: zorluk seviyesi, soru tipi) zaman alıcı ve karmaşık olabilmesi.  Zaman planının sınırlı olması, bazı analizlerin (örn: dilbilgisi yapısı derin analizi) yüzeysel kalmasına neden olması.  Geliştirilen ilk modelin tahmin doğruluk oranının hedeflenenin altında kalma riski.	MEB'in (Milli Eğitim Bakanlığı) LGS soru formatını veya müfredat ağırlıklarını aniden değiştirmesi.  Modelin, geçmiş verilere aşırı odaklanıp (overfitting) yeni trendleri kaçırarak hatalı tahminler üretmesi.  Zaman yönetimi ve ekip içi koordinasyonun aksaması durumunda projenin gecikmesi.  Veri toplama veya analiz aşamasında beklenmedik teknik sorunlar (örn: veri uyumsuzluğu).